Abstract

Separability is a concept that is very difficult to define, and yet much of our scientific method is implicitly based upon the assumption that systems can sensibly be reduced to a set of interacting components. This paper examines the notion of separability in the creation of bi-ambiguous compounds that is based upon the CHSH and CH inequalities. It reports results of an experiment showing that violations of the CHSH and CH inequality can occur in human conceptual combination.

1 Separability and its Representation

Separability is a concept that is very difficult to define, and yet much of our scientific method is implicitly based upon the assumption that systems can sensibly be reduced to a set of interacting components. This assumption has been captured in the hidden variable models of physics, which make explicit the separability of a system in terms of probability distributions. Thus, if the two components of a system are somehow ‘distant’ then it makes sense to assume that actions performed upon one region will not effect the results that are obtained in the other region. A general scenario representing this case is illustrated in figure 1.

Figure 1: A general scenario illustrating the separability of a system consisting of two distant components.

Here, a system consisting of two components is frequently considered separable due to the fact that the components are in some sense distant (generally with respect to a physical space, although any space possessing a metric can be considered to have distant components). Thus, making a choice in region $A$ to measure some characteristic $c_A$ of the system is deemed not to effect the results that will be obtained when the characteristic $c_B$ is measured in region $B$.

This separability can be represented by the assumption that random variables representing the characteristics in one region (call them $A$ say) do not interact with those in the second region ($B$). This implies that, accounting for any set of hidden variables $\lambda$ with the probability distribution $\rho(\lambda)$, we find that the joint probability of the two random variables is

$$\Pr(A, B) = \int d\lambda \rho(\lambda) \Pr(A|\lambda) \Pr(B|\lambda), \tag{1}$$

implying that we can make probability statements about the whole system by performing a multiplication operation upon probability statements that arise for each of its components.

Immediately, we can see that a system which violates (1) is in some sense non-separable. However, constructing experimental tests of this behaviour can be a very difficult process (Laloë 2001). Indeed, while Bell proposed his now infamous inequality using an analogue of this equation in 1964, it took almost 20 years before experimental tests were performed that satisfied the bulk of the physics community (and loopholes allowing local realistic interpretations of quantum theory still exist although the models are becoming less and less plausible).

Despite its well-known history in modern physics, statements of the general form represented by equation (1) are not specific to Quantum Theory (QT). They are widely used in science. However, there are many reasons to believe that they do not generally apply to all systems (Aerts 2000; Aerts et al. 2000; Iqbal and Cheon 2007). In this paper we shall discuss one particular case of non-separable behaviour that arises in cognition.

2 When are Concepts Non-separable?

This work will extend previous work showing that concepts can behave non-separably (Gabora, Rosch, and Aerts 2008; Aerts 2009). It will proceed by looking at the way in which humans process novel concept combinations. It is very often the case that when humans are presented with unfamiliar word pairings they generate a meaning for them that is
highly imaginative. Indeed, they can frequently generate emergent associations for the new combined concept (i.e. associations that are not present in either of the individual concepts). However, it is not the case that ‘anything goes’. There is very often an agreement between subjects that a meaning ascribed to such a concept combination is sensible, even if it is not the individual’s favourite definition of an unfamiliar word pair.

As an illustrative example, consider the concept combination “pet human”. In word association experiments, human subjects quite readily ascribe the associates “slave”, “robot” or “child” to this combination. However, none of these associates are produced in relation to “pet”, or “human” when they are individually presented as a cue in a free association experiment (Nelson, McEvoy, and Schrieber 2004). So it seems that these new associates are emergent in the sense that they arise from the concept combination as a whole, and cannot be recovered from its constituent concepts.

There has been growing speculation in the literature that concept combinations may behave like quantum entangled particles (Aerts and Gabora 2005; Gabora, Rosch, and Aerts 2008; Aerts 2009; Bruza et al. 2010). Indeed, QT provides a very natural model of non-separability and we shall make use of this account in what follows, where we shall make a very natural model of non-separability and we shall make the hypothesis that a full account of conceptual combination cannot be provided by reductive modelling. Thus, we shall explore the idea that conceptual combination cannot be fully understood in terms of the constituent words in the combination; a concept combination is not simply a “sum of its parts”.

In this paper we shall discuss a recent experiment used to test these ideas. We have run this experiment for the novel concept combinations presented in table 1, which take ambiguous words (which have at least two different senses, each of which is representing a different concept) and combine them in novel ways. We can represent these ambiguous words using a variable A, which ranges over \{a, a’\} corresponding to the two underlying senses, and similarly B ranges over \{b, b’\}.

This notation is readily linked to the notion of separability that was introduced in equation (1). In word association experiments, subjects are cued by a priming word that we shall represent using \(\lambda\). For the case of bi-ambiguous concepts, it is interesting to consider the way in which we can bias the meaning that a subject attributes to a word through a careful choice of cues. Thus, for each ambiguous word we should be able to identify a cue that triggers one of the two senses. In order to illustrate these ideas we shall consider the concept combination “boxer bat”. Here, “boxer” has an animal sense, and a sporting sense, as does “bat”. Thus, representing “boxer” with the variable A ranging over two senses \{a = animal, a’ = sporting\}, we should be able to find two cues that will engage these two senses: \{\lambda_1 = dog, \lambda_2 = fighter\} say. Similarly for “bat”, we could take the two senses to be \{b = animal, b’ = sporting\} and the respective cues: \{\lambda_3 = vampire, \lambda_4 = ball\}.

Here, we are modelling the primes as a random variable \(\lambda\) ranging over \{\lambda_1, \lambda_2, \lambda_3, \lambda_4\}. In this probabilistic setting, we can re-formulate (1) as

\[
\Pr(A, B) = \sum_{1 \leq i \leq 4} \Pr(A|\lambda_i) \Pr(B|\lambda_i) \Pr(\lambda_i). \tag{2}
\]

It is interesting to consider the underlying assumptions in obtaining this relation. Essentially we have started from the underlying assumption that the joint probability of \(A\) and \(B\) is factorizable:

\[
\Pr(A, B|\lambda) = \Pr(A|\lambda) \Pr(B|\lambda), \tag{3}
\]

Thus, separability has been reduced to factorizability. This is an equivalence that is not necessarily correct, a problem that is discussed in more detail in (Bruza, Iqbal, and Kitto 2010). We can then apply Bayes’ rule, to rewrite the joint probability as:

\[
\Pr(A, B, \lambda) = \Pr(A|\lambda) \Pr(B|\lambda) \Pr(\lambda) \tag{4}
\]

and finally, assuming the law of total probability, we obtain (2).

This paper will test these assumptions through the investigation of a concept combination experiment based upon a Bell-type inequality, in this case the Clauser–Horne–Shimony–Holt (CHSH) inequality. These inequalities arise when a similar set of separability assumptions are made in Quantum Theory (QT) (Laloë 2001; Greenstein and Zajonc 1997). Entangled quantum systems violate this inequality and hence do not conform to one of the assumptions made in obtaining it. This result shows that entangled quantum systems exhibit non-separable effects, although the precise form of that non-separability is yet to be well defined in the quantum foundations literature.

Concept combination is still largely an open question within cognitive science, so evidence of similar quantum-like effects would have significant ramifications. It would suggest that a particular set of classical models will not provide complete models of the processes involved in human reasoning, which should in turn influence current directions in Artificial Intelligence (AI).

3 Entangling Words and Meaning

In (Bruza et al. 2008; 2009) we proposed a set of experimental designs that might be used to test for the quantum-like behaviour of words in the human mental lexicon. This paper builds upon that work, through the provision of the specific set of compounds reported in table 1, and the discussion of a specific experimental protocol and preliminary results. This section shall motivate the decisions that were made in designing the experiment reported in section 4.

The CHSH inequality provides an experimental test for distinguishing between local hidden variables theories (as are exemplified by equation (1) above) and entangled (i.e. non-separable) systems. The quantum scenario is illustrated in figure 2. In the basic scenario, a source \(S\) emits two entangled photons, one travels left through the polariser at \(c_A\), the other right through the polariser at \(c_B\). The photons can reflect from the polariser, or transmit through it, and the state describing the system becomes more complex again representing the different likelihoods of this occurring. Finally,
Figure 2: An experimental scenario testing for the non-separability of an entangled system of polarised photons. A source emits two entangled photons that travel to polarisers at $c_A$ and $c_B$. In each of the regions $A$ and $B$, either detector $D_0$ or $D_1$ clicks, and this is recorded at a coincidence counter.

two detectors in this system ‘click’, one on the left side, and one on the right. Coincidence is measured in this scenario, with $|11\rangle$ representing a situation where the two detectors requiring transmission through the polariser click, and so on for the other states. Finally, the orientation of these polarisers can be changed, and this leads to a different proportion of photons being transmitted or reflected.

The results of this experiment are used to calculate expectation values for the four available combinations of two different polariser settings, $a, a', b, b'$:

$$E(i, j) = \frac{N_{11} + N_{00} - N_{10} - N_{01}}{N_{11} + N_{00} + N_{10} + N_{01}}$$

where $i \in \{a, a'\}, j \in \{b, b'\}$. (5)

If the two different sides of this experiment can be considered separately, then the expectation values for this experimental scenario will satisfy the CHSH inequality:

$$-2 \leq E(a, b) - E(a, b') + E(a', b) + E(a', b') \leq 2$$

which provides us with a numerical test for the separability (or not) of a quantum system. If the system can be considered separable then the CHSH inequality will be satisfied. This then means that it is possible to consider the parts of the system in isolation. Quantum systems can violate equation (6) and hence should not be treated in this manner.

In modifying this scenario such that it can be applied to human cognitive states a number of decisions had to be made.

Firstly, notice that in constructing the expectation value, we require two polariser settings, one for each component of the experiment. Furthermore, these values count the co-occurrences of two results at each side of the experiment. That is, continuing with our “boxer bat” example above, questions such as ‘how often does the animal sense of both words arise?’, or ‘how often does a animal sense of “boxer” and a sport sense of “bat” arise? ‘ are asked. However, a complicating factor occurs in the attempt to consider what is meant by the concept of the relative ‘orientations’ between cues in semantic spaces. In the quantum CHSH experiments, the concept of coincidence events and polarisers is straightforward to define: it is the count of the number of co-occurrences of a given event defined by the orientation of two polarisers. It is clear that we are talking about the same type of event for the quantum scenario. That is, we can compare the orientation of the two polarisers, simply by comparing the angles at which they are placed, and we are clearly talking about the same kind of measurement. However, for the case of words, we do not know how to do this; are we comparing the same objects? That is, are word senses all in the same space (like polarisation), or must they be considered differently? In what follows we have attempted to choose concept pairs that have overlapping senses such as the animal and sport senses shared by “boxer” and “bat” (see table 1 for the full list). This choice, while it does not allow for the full range of ‘angles’ used in the quantum CHSH experiment does ensure that the words are at least comparable.

Returning to the example of the compound “boxer bat”, it is possible to use the four primes defined above in a set of four experiments:

1. $(a, b) = (\text{fighter}, \text{ball}) \sim (\text{sport}, \text{sport})$ senses
2. $(a, b') = (\text{fighter}, \text{vampire}) \sim (\text{sport}, \text{animal})$ senses
3. $(a', b) = (\text{dog}, \text{ball}) \sim (\text{animal}, \text{sport})$ senses
4. $(a', b') = (\text{dog}, \text{vampire}) \sim (\text{animal}, \text{animal})$ senses

A “boxer bat” has a number of possible interpretations. It might be construed as “a small furry black animal with boxing gloves on”, or perhaps it could be a “baseball bat a boxer dog plays with”. In each of these interpretations we see that a different sense of the component concepts has been taken. A subject who decides upon the first interpretation will have an activation of the animal sense of ‘bat’ and the sport sense of ‘boxer’, while a second subject will have chosen sport sense of ‘bat’ and the animal sense of ‘boxer’. We shall choose to represent these activation states (of the different concept senses) with reference to the experimental setting (in a manner very similar to standard quantum mechanics). Thus, if a subject is cued with a word that has a sport sense, and they return a sport interpretation of the bi-ambiguous concept then provided, then they will be deemed to have been in an activated, or $|1\rangle$ state, if not, then they will be in a state of non-activation $|0\rangle$.

We expect that there will be a statistical norm that will provide a ‘natural’, or unbiased interpretation of a concept combination; it is the most likely meaning that an unbiased subject will attribute to the combination. However, if a subject is exposed to two cues that bias the senses ascribed to the component concepts then we may find that we can shift the ‘natural’ interpretation that they would ascribe to the compound. This would start to illustrate the non-separable nature of concept combinations in a statistically meaningful manner. If a strong biasing could be generated then this could potentially lead to violations of the separability assumptions discussed above.

How should we count detection events for concepts? In this experiment we have chosen to count events as follows; if a subject returns an interpretation for a concept that agrees with the one they were cued with then a $|1\rangle$ will be recorded, if they disagree, then a $|0\rangle$ will be deemed to have occurred. So, if a subject sees the two cues “fighter” and “vampire”,
and then deems that a “boxer bat” is “a small furry black animal with boxing gloves on” then they will have scored a |11⟩ and the $N_{11}$ count will be increased, while if they deemed that it was a “baseball bat a boxer dog plays with” then they will have scored a |00⟩, with a corresponding increase in $N_{00}$.

4 Experimental Design

Table 1 lists all compounds that were tested, along with the words used to define polariser settings (which are given by the senses listed alongside in italics). Each compound represents a new experiment, while the senses, which apply to each word in the compound separately, could be regarded as hidden variables. Cues are taken to correspond to polarisers oriented in some particular direction in this scenario, and they are somewhat more restrictive than those used in a standard Bell-type arrangement, as they are oriented in the same direction as the hidden variables. This is a very special scenario, not generally used in the more standard quantum tests as in this case there is no corresponding value for the ‘spin’ of an unmeasured quantum particle.

For the purpose of clarifying how the above experimental scenario was implemented, we shall first walk through one task for the concept combination “boxer bat”.

Participants completed an online experiment in which they were asked to provide an interpretation for twelve compounds (e.g., “boxer bat”). Each compound was seen only once by a participant. For groups 1-8, each compound was preceded by a similarity rating task (see figure 3), in which participants rated the similarity between two pairs of words (e.g., “dog” and “boxer”, “vampire” and “bat”) on a 7 point scale (low similarity to high similarity). For the purpose of clarifying how the above experimental scenario was implemented, we shall first walk through one task for the concept combination “boxer bat”.

Figure 3: The first similarity rating task. This attempts to bias the sense chosen by the subjects when they are presented with the compound. It effectively biases, or ‘polarises’ the meanings of the concepts.

This was done to simultaneously prime the words of the compound into two of the four possible combinations of senses (here, the animal-animal case is defined from table 1 as s1-s1). The groups were designed so that over the 12 compounds, each participant received the 4 possible cues pairs (s1-s1, s1-s2, s2-s1, s2-s2) 3 times. As an example, for boxer bat, group 1 received the cue pair s1-s1, group 2 received the cue pair s2-s1, group 3 received the cue pair s2-s1, and group 4 received the cue pair s2-s2. For bank log, the cue pairs for groups 1-4 were respectively s1-s2, s2-s1, s2-s2, s1-s1, for apple chip s2-s1, s2-s2, s1-s1, s1-s2, and so on through all of the compounds. Groups 5-8 counter balanced this design by reversing the ordering. For example, for ‘boxer bat’, groups 5-8 respectively received the cue pairs: s2-s2, s2-s1, s1-s2, s1-s1 (the reverse of groups 1-4) and so on. Groups 9 and 10 received neutral primes in which they categorised two numbers (e.g., 35 and 72) as odd or even. These were the baseline groups in which the compounds received no priming, but rather the senses of the two words were selected solely by the participants.

In the second stage, participants are asked to provide an interpretation for a novel compound. This compound is formed from taking the two words appearing last in the similarity test during the first stage (see figure 3). This stage performs the role of a quantum measurement, in providing an interpretation for the compound participants must ‘collapse’ each of the ambiguous words in the compound to one particular sense.

Figure 4: The second task, where subjects must provide an interpretation for a novel compound.

Finally, after interpreting the compound, participants were asked to clarify which sense they chose for each word. For boxer participants could select: (A) An animal, (B) A fighter, or (C) Other (which they were asked to specify). This removes any potential ambiguity surrounding which word sense was applied in constructing the new concept represented by the novel compound.

Figure 5: The third task, where subjects must define the sense that they used in the definition of the compound for each bi-ambiguous word.

Some further comments about this experimental design are appropriate. First, it is important to note that this experiment is, in some sense, backwards. The ambiguous words are sent through their polarising phase before they are actually considered as a pair, or entangled in the standard quantum scenario. This difference with the standard CHSH sce-
Table 1: The compounds chosen for this experiment. Each has the same two senses or interpretations, s1 and s2. These compounds are primed by two of four possible cues, each of which biases the compounds towards a certain sense (listed in the table).

<table>
<thead>
<tr>
<th>Compound</th>
<th>sense 1 (s1)</th>
<th>sense 2 (s2)</th>
<th>a (s1)</th>
<th>b (s1)</th>
<th>a' (s2)</th>
<th>b' (s2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>boxer bat</td>
<td>sport</td>
<td>animal</td>
<td>fighter</td>
<td>ball</td>
<td>dog</td>
<td>vampire</td>
</tr>
<tr>
<td>bank log</td>
<td>natural</td>
<td>financial</td>
<td>river</td>
<td>cabin</td>
<td>money</td>
<td>journal</td>
</tr>
<tr>
<td>apple chip</td>
<td>food</td>
<td>computer</td>
<td>banana</td>
<td>potato</td>
<td>computer</td>
<td>circuit</td>
</tr>
<tr>
<td>stock tick</td>
<td>financial</td>
<td>animal</td>
<td>shares</td>
<td>mark</td>
<td>cow</td>
<td>flea</td>
</tr>
<tr>
<td>seal pack</td>
<td>container</td>
<td>animal</td>
<td>envelop</td>
<td>suitcase</td>
<td>walrus</td>
<td>leader</td>
</tr>
<tr>
<td>spring plant</td>
<td>natural</td>
<td>artefact</td>
<td>summer</td>
<td>seed</td>
<td>coal</td>
<td>factory</td>
</tr>
<tr>
<td>poker spade</td>
<td>cards</td>
<td>implement</td>
<td>card</td>
<td>ace</td>
<td>fire</td>
<td>shovel</td>
</tr>
<tr>
<td>slug duck</td>
<td>body action</td>
<td>animal</td>
<td>punch</td>
<td>dodge</td>
<td>snail</td>
<td>quack</td>
</tr>
<tr>
<td>club bar</td>
<td>place</td>
<td>artefact</td>
<td>member</td>
<td>pub</td>
<td>golf</td>
<td>handle</td>
</tr>
<tr>
<td>web bug</td>
<td>insect</td>
<td>computer</td>
<td>cob</td>
<td>beetle</td>
<td>Internet</td>
<td>computer</td>
</tr>
<tr>
<td>table file</td>
<td>artefact</td>
<td>record</td>
<td>chair</td>
<td>nail</td>
<td>chart</td>
<td>folder</td>
</tr>
<tr>
<td>match bowl</td>
<td>sport</td>
<td>artefact</td>
<td>contest</td>
<td>throw</td>
<td>flame</td>
<td>dish</td>
</tr>
</tbody>
</table>

Table 2: The sets of coincidence data and the expectation value (calculated using (5) obtained from these coincidences for the four different “polariser settings” used in the ‘boxer bat’ experiment. Here, a represents the cue fighter, b: ball, a’: dog, and b’: vampire.

<table>
<thead>
<tr>
<th>experiment</th>
<th>N_{11}</th>
<th>N_{10}</th>
<th>N_{01}</th>
<th>N_{00}</th>
<th>E(x, y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a, b)</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>-0.2</td>
</tr>
<tr>
<td>(a, b')</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>-0.67</td>
</tr>
<tr>
<td>(a', b)</td>
<td>1</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>-0.56</td>
</tr>
<tr>
<td>(a', b')</td>
<td>3</td>
<td>0</td>
<td>5</td>
<td>1</td>
<td>0.17</td>
</tr>
</tbody>
</table>

5 Results and Discussion

Continuing with our consideration of ‘boxer bat’, we shall now estimate the expectation values for the results obtained and use these to calculate the expectation value. So, for experiment (a, b), we have the primes set to (sport, sport) senses, hence any participants that returned a meaning for the compound that involved them identifying both ‘boxer’ and ‘bat’ in the sport sense will result in a correlation (N_{11}). Table 2 lists the results obtained. Note that a fairly high spread of counts was obtained for this experiment, with (in this case) 34 counts arising that were deemed to be legitimate counts (i.e. not listed as Other in the third task). This is a reasonably low count rate, generally over 40 legitimate counts were recorded — see the data file at http://www.quantuminteraction.org/conceptCombinationExpts/data for further details (use data/data for login/password). Obviously this is still too low a count to be significant, however, as this was a pilot study the numbers of participants was not increased. A larger, more controlled study is currently underway.

Now we can use these data to calculate the CHSH inequality for the combination ‘boxer bat’:

\[
CHSH = E(a, b) - E(a', b') + E(a', b) + E(a', b') = 0.08
\]

(7)

In this case, we see that the value extracted, \(-2 \leq CHSH \leq 2\); a violation has not occurred. However, an ambiguity presents itself in carrying out this analysis. Choices must be made as to which experiment receives which label, and these choices can have a profound effect. Thus, in the current scenario, changing around the primed vs the unprimed cues results in an outcome of \(CHSH = -0.14\) (which is still within the bounds of the inequality). This asymmetry is not particularly surprising upon some thought, although it is not frequently talked about in the physics literature which tends to choose experimental scenarios that guarantee maximal violations. The antisymmetry of equation (6) implies that the way in which experiments are labelled can affect the final inequality, indeed, the same set of experiments can change from yielding a violation to not doing so, simply by reordering the labels. This is not important to the main result; if a violation can be obtained from the data then the system is deemed to have violated the CHSH inequality. In what follows, we list the maximal violations obtained for the experimental scenarios that were created.

We have performed the analysis described above for boxer bat with every concept combination listed in table 1, and have obtained the results listed in table 3. Because of the webbased nature of these experiments, participants were not always assigned equally to each group of analyser settings; different numbers of participants were exposed to each setting, but in every case over 40 participants were exposed to each general task (i.e. to each biambiguous compound) when the neutral groups are considered too. For more details of the actual calculations performed...
in obtaining the following data readers are encouraged to consult the raw data file at http://www.quantuminteraction.org/conceptCombinationExpts/data together with the analysis files. Use the login/password pair data/data to view the files.

<table>
<thead>
<tr>
<th>Compound</th>
<th>CHSH</th>
<th>CH</th>
</tr>
</thead>
<tbody>
<tr>
<td>boxer bat</td>
<td>0.08</td>
<td>-0.67</td>
</tr>
<tr>
<td>bank log</td>
<td>1.92</td>
<td>-0.15</td>
</tr>
<tr>
<td>apple chip</td>
<td>2</td>
<td>-0.09</td>
</tr>
<tr>
<td>stock tick</td>
<td>2.14</td>
<td>0.01</td>
</tr>
<tr>
<td>seal pack</td>
<td>1.94</td>
<td>-0.03</td>
</tr>
<tr>
<td>spring plant</td>
<td>1.94</td>
<td>-0.07</td>
</tr>
<tr>
<td>poker spade</td>
<td>2</td>
<td>-0.26</td>
</tr>
<tr>
<td>slug duck</td>
<td>2.04</td>
<td>-0.46</td>
</tr>
<tr>
<td>club bar</td>
<td>2.04</td>
<td>0</td>
</tr>
<tr>
<td>web bug</td>
<td>1.83</td>
<td>-0.11</td>
</tr>
<tr>
<td>table file</td>
<td>-0.09</td>
<td>-0.57</td>
</tr>
<tr>
<td>match bowl</td>
<td>2.02</td>
<td>0.17</td>
</tr>
</tbody>
</table>

Table 3: Expectation values for the experimental settings given in table 1, along with the value CHSH calculated according to equation (7), and CH results calculated according to equation (8). The full dataset used to obtain these results is available at http://www.quantuminteraction.org/conceptCombinationExpts/data together with the analysis files. Use the login/password pair data/data to view the files.

Table 3 shows that a violation of the CHSH inequality was found for four bi-ambiguous word pairs (“stock tick”, “slug duck”, “club bar” and “match bowl”). Thus, we have found some preliminary indications that concept combinations can indeed behave non-separably in some situations. However, we must consider this result in more detail before we can proclaim success.

Violators of the CHSH inequality do not appear to follow any strong trends. There is a tendency for violations to be recorded in cases where both ambiguous words are interpreted with the same sense (typified by a high expectation value in the \((a,b)\) and \((a',b')\) cue scenarios), but this is not always the case. Certain behaviour must be exhibited by the anticorrelated polarisation cases as well. Examining equation (6) suggests a number of plausible violation scenarios constructed around maximal expectation values, but none of our experiments yielded such results. This problem causes us to pause: is the failure to get a strong violation a function of the cueing procedure, the experimental procedure, or the human mental lexicon? At present we are unable to answer this question. However, we can make a number of observations.

First, we must note that our experiment is prone to a detection loophole like problem (Laloë 2001; Ballentine 1998). This loophole arises from cases where not all detection events are recorded. The possibility of responding to the test depicted in figure 5 with an ‘Other’ classification, opens the current analysis up to this problem as we have for the present ignored these responses. This becomes more problematic when we consider the relative number of ‘C’ options recorded for the violators (which was high in comparison to those cases where no violation was recorded).

However, we can undertake a proper consideration of this scenario through the adoption of the Clauser–Horne (CH) inequality (Clauser and Horne 1974). This inequality uses probability of coincidences, instead of the expectation values used in (6). Thus, \(p(i,j)\) corresponds to the probability that the experiment \((i,j)\) gives the outcome 1, 1: \(p(i,j) = \frac{N_{11}}{N_{11} + N_{00}}\). The CH inequality adds two new experimental arrangements to the CHSH inequality, representing independence-style assumptions corresponding to the probability \(p(i)\) that a single wing of the experiment (in region \(A\) say) gives the outcome 1 when its analyser is set to experiment \((i)\).

\[
-1 \leq p(a,b) - p(a,b') + p(a',b) + p(a',b') - p(a') - p(b) \leq 0
\]

Thus, this equation takes into account the null detection events on side \(A\) and \(B\) where a result is not recorded. An analysis using the CH inequality has been performed and the results are reported under the CH column in table 3. We see that two of the original CHSH violations have been lost due to detection loophole problems, but two violations remain under this analysis (“stock tick” and “match bowl”).

It is important however, to keep in mind that these results are only preliminary. Our experiment effectively consisted of 12 separate CHSH experiments and sample sizes are not yet large enough for these results to be considered robust. The choice to perform a number of different scenarios was made in an attempt to cover a wide set of possibilities. The difficulty in mapping the CHSH inequality directly into the case of conceptual combination made it necessary to keep this experiment as broad as possible in order to make the chance of achieving a violation as high as possible. However, a larger data set will be required to strengthen these results, and work is progressing upon this.

We feel that with a more robust sample size, these techniques might yield stronger results. After all, for every concept combination created, the state of the human mind generating it is tested, which makes the case of completely undetected states less obviously relevant. More difficult to resolve is the problem of subjects creating multiple interpretations and then reporting only one. It seems likely that some subjects are rejecting their first associations as unlikely or wrong, choosing instead to report upon a subsequent interpretation. In order to limit such responses it seems likely that we shall have to shift from web-based experiments to more tightly supervised scenarios in the future, however, it might also be that a change in cueing procedure could help in strengthening these results.

Priming is an on-going problem with these experiments. The procedure utilised in this experiment was chosen with the expectation that the cognitive task of rating the similarity of the presented words was not too heavy, but that it would cause subjects to think about the ambiguous words forming the compound in a certain context. The presentation of both sets of primes together with their respective ambiguous word on the same page was expected to stop bias between the two ‘wings’ of the experiment. However, there is reason
to believe that this task was not as straightforward as was expected.

For example, there is an unresolved question over the effectiveness of the priming method used in this experiment. The similarity decision was expected to isolate priming to each individual word of the compound (e.g., “ball” primed the sports sense of “bat” and “vampire” primed the animal sense of “bat”). If the primes did work then one would expect that sport-sport cues, \((a, b)\), would push “boxer” and “bat” to their sport senses, and the animal-animal cues, \((a', b')\), would push the interpretation given to the animal interpretation of the bi-ambiguous compound. However, in many cases the prime did not shift participants in the expected direction.

For example, we found that on average 7 \(ab\) interpretations were given for \((a, b)\) cues and 6.17 were given for \((a', b')\) cues. While participants gave less \(ab\) interpretations when primed with \((a', b')\) senses (as would be expected if the priming was working), a between groups t-test showed no significant difference between the two conditions, \(t(22) = 0.53, p = .60\). However since observed power was also very low (.08), more compounds may be needed to establish whether or not the priming was effective. Similarly, an average of 3 \(a'b'\) interpretations were recorded for their equivalent \((a', b')\) cues and 3.57 recorded for \((a, b)\) cues. As this result is occurring in the reverse direction, it did not show the expected effect of priming (i.e. more results were anti-correlated with their cues in this case).

Thus, there is some reason to believe that the priming procedure was not particularly effective in this experiment. Rather than creating the intended ‘entangled’ state of mind, it may have been distracting subjects from the main task. As a specific example of this phenomenon (where the priming did not seem to work) we found that participants rated “fighter” and “boxer” as highly similar \((M = 5.35)\) and “dog” and “boxer” as very dissimilar \((M = 1.95)\). However, if participants were thinking of “boxer” in the dog sense then they should have rated “dog” and “boxer” as highly similar, and at least as similar as fighter to boxer (as boxer dog is a dog, just as much if not more so, as a human boxer is a fighter).

A between groups t-test revealed that the similarity ratings between these pairs was significantly different, \(t(41) = 9.33, p < .001\). In fact, a number of participants gave informal feedback saying that they found the similarity rating task to be more difficult than the primary task, and many appeared to be somewhat bemused by this stage of the experiment, which suggests that a task expected to be relatively simple created a heavy cognitive load. This may have led to unexpected outcomes in the experimental results.

In the larger trial that is currently being implemented, subjects will be asked to consider a set of primes and compounds that are not all legal English words. In the priming stage they will be asked to decide whether the presented words are legal ‘words’ or ‘non-words’, before being asked to provide interpretations for novel compounds. This task is expected to be cognitively more straightforward, while still forcing subjects to consider the primes and the compounds together (thus ‘entangling’ them together non-separably). We expect that a more effective priming procedure would lead to larger and more consistent violations.

Quantum-like entanglement has not been shown by this experiment. Indeed, even in standard quantum theory the CHSH experiments do not illustrate that a quantum system must be considered as entangled. Rather, such experiments show that a local hidden variables model (as represented by the separability assumption in \((1)\)) of these systems is inappropriate. In the current case, we have a preliminary indication that concept combinations should be considered as non-separable, but does this imply entanglement? Perhaps this question is ill-phrased. It is clear that many non-separable scenarios can be best represented using a formalism that includes entangled states, similarly, it is clear that CHSH tests can often point to those situations where such formalism is useful. In adapting the quantum formalism to the description of macroscopic ‘classical’ systems it is perhaps only necessary to show a similar set of conditions apply, although this will no doubt be considered inadequate by those who question such extensions of the quantum formalism. At present the use of the term ‘entangled’ is perhaps more of a question of utility (and indeed its use has been minimised in this paper). With experiments exhibiting quantum-like interference between concepts we could perhaps start to answer such criticism, as this would be another effect that is not well modelled by a classical probability structure. Searching for such behaviour will be an area for future work.

In summary, it must be emphasised that this data can only be treated as preliminary. As participants could only be exposed to one choice of ‘polariser settings’ each, we will require a much larger sample before we can begin to consider this work as statistically significant. Now that we have some indications as to how these experiments might be successfully run we are in a much firmer position to run a larger experiment.

6 Conclusions

It seems that the context in which a word appears can profoundly influence the interpretation attributed to it. This work has taken a very particular scenario, considering cases where humans are confronted with novel biambiguous compounds representing mutually exclusive concept combinations. This means that they are forced to provide a novel interpretation for the new compound, but it seems that the interpretation provided can be influenced by the context in which the novel compound appears. This is not a particularly surprising outcome. Indeed, psychology and linguistics both recognise the inherently nonseparable nature of concept combination (Gagne and Shoben 1997; Gärdenfors 2000). However, this work has illustrated that such behaviour can be both modelled theoretically, and probed experimentally.

Thus, while the formulation of this experiment is quite restrictive and these results might therefore appear to be rather specialised, we do not think that this is likely to be the case. Consider for example the concept of a ‘chair’. A different interpretation of this concept arises from almost every context in which it appears, thus an ‘office chair’ is quite different from a ‘kitchen chair’, and novel combinations like ‘ocean
Indeed, while the notion of a chair might seem very straight-forward, this is still a highly modifiable concept, and the creativity of humans is exceptionally adept at providing a set of reasonable interpretations for such new concepts. Humans are frequently confronted with situations where they must provide new creative interpretations for concept combinations, and yet we have very few theoretical constructs modelling the way in which they achieve such interpretations. This paper provides one set of tools for modelling this nonseparable behaviour, but much more will be required for a complete theoretical model of concept combination.

For example, the Bell-like tests reported in this paper take their inspiration directly from quantum theory, but we could well envisage a set of tests for non-separable behaviour that is derived directly from psychology. What forms would such tests take? At present we cannot even speculate, but it is likely that the quantum formalism is not the only one that can be used in this scenario.

Further into the future, we might envisage a ‘concept combiner’ that could take two concepts and provide a set of statistically weighted possible meanings which correlate with human suggestions for novel concepts. The conceptual, theoretical and implementational difficulties that beset such a programme of research are large, but the challenge is likely to be surmountable. How much of such a tool will owe its origins to a quantum-like model remains to be seen, but this paper gives some reason to believe that a non-separable model will be necessary. Concepts should not be modelled in isolation, but in the context in which they occur.

While concept combination undoubtedly presents a major problem for artificial systems, we can see a way in which to proceed. It is hoped that this initial set of results will inspire others to consider working in this particularly interesting field.

References


1The authors are indebted to Doug Nelson for this example.